VIETNAM ACADEMY OF SCIENCE AND TECHNOLOGY
INSTITUTE OF MATHEMATICS

NGUYEN THAI AN

Variational Analysis and Some Special Optimization Problems

Speciality: Applied Mathematics
Speciality code: 62 46 01 12

SUMMARY OF

DOCTORAL DISSERTATION IN MATHEMATICS

HANOI - 2016

The dissertation was written on the basis of the author's research works carried out at the Institute of Mathematics, Vietnam Academy of Science and Technology

Supervisors:

1. **Prof. Dr. Hab. Nguyen Dong Yen**
2. **Assoc. Prof. Nguyen Mau Nam**

First referee: ....
....

Second referee: ....
....

Third referee: ....
....

To be defended at the Jury of the Institute of Mathematics, Vietnam Academy of Science and Technology:

....
....
on .... 2016, at .... o'clock ....

The dissertation is publicly available at:

- The National Library of Vietnam
- The Library of Institute of Mathematics

Introduction

Optimization techniques usually require differentiability of the function involved, while nondifferentiable structures appear frequently and naturally in many mathematical models. Motivated by applications to optimization problems with nondifferentiable data, *variational analysis* has been developed to study generalized differentiability properties of functions, and set-valued mappings without imposing the smoothness of the data.

Facility location, also known as location analysis, is a branch of operations research and computational geometry that concerns with mathematical modelings and solution methods for problems of finding the right site of a set of facilities in a given space in order to supply some service to a set of demands/customers. Depending on specific applications, location models are very different in their objective functions, the distance metric applied, the number and size of the facilities to locate; see, e.g., Z. Drezner and H. Hamacher, *Facility Location: Applications and Theory*, (Springer, Berlin, 2002) and R. Z. Farahani and M. Hekmatfar, *Facility Location: Concepts, Models, Algorithms and Case Studies*, (Physica-Verlag Heidelberg, 2009), and the references therein.

The origin of location theory can be traced back as far as to the 17th century when P. de Fermat (1601-1665) formulated the problem of *finding a fourth point such that the sum of its distances to the three given points in the plane is minimal.* This celebrated problem was then solved by E. Torricelli (1608-1647). At the beginning of the 18th century, A. Weber incorporated weights, and was able to treat facility location problems with more than 3 points as follows

$$\min\Big\{\sum_{i=1}^{m}\alpha_i\|x-a_i\|:\ x\in I\!R^n\Big\},$$

where $\alpha_i > 0$ for $i = 1, \ldots, m$ are given weights and the vectors $a_i \in I\!R^n$ for $i = 1 \ldots m$ are given demand points.

The first numerical algorithm for solving the Fermat-Torricelli problem was introduced by E. Weiszfeld (1937). As pointed out by H. W. Kuhn (1973), the Weiszfeld algorithm may fail to converge when the iterative sequence enters the set of demand points. The assumptions guaranteeing the convergence of the Weiszfeld algorithm along with a proof of the convergence theorem were given by Kuhn. Generalized versions of the Fermat-Torricelli problem and several new algorithms have been introduced to solve generalized Fermat-Torricelli problems as well as to improve the Weiszfeld algorithm. The Fermat-Torricelli problem has also been revisited several times from different viewpoints.

The Fermat-Torricelli/Weber problem on the plane with some negative weights was first introduced and solved in the triangle case by L.-N Tellier (1985) and then generalized by Z. Drezner and G. O. Wesolowsky (1990) with the following formulation in $\mathbb{R}^2$:

$$\min\Big\{\sum_{i=1}^{p}\alpha_i\|x-a_i\|-\sum_{j=1}^{q}\beta_j\|x-b_j\| \,:\, x\in\mathbb{R}^2\Big\}, \tag{1}$$

where α_i for $i=1,\ldots,p$ and β_j for $j=1,\ldots,q$ are positive numbers; the vectors $a_i\in\mathbb{R}^2$ for $i=1,\ldots,p$ and $b_j\in\mathbb{R}^2$ for $j=1,\ldots,q$ are given demand points. According to Z. Drezner and G. O. Wesolowsky, a negative weight for a demand point means that the cost is increased as the facility approaches that demand point. One can view demand points as attracting or repelling the facility, and the optimal location as the one that balances the forces. Since the problem is nonconvex in general, traditional solution methods of convex optimization widely used in the previous convex versions of the Fermat-Torricelli problem, are no longer applicable to this case. The first numerical algorithm for solving this nonconvex problem which is based on the outer-approximation procedure from global optimization was given by P.-C. Chen, P. Hansen, B. Jaumard, and H. Tuy (1992).

The *smallest enclosing circle problem* can be stated as follows: *Given a finite set of points in the plane, find the circle of smallest radius that encloses all of the points.* It was introduced in the 19th century by the English mathematician J. J. Sylvester (1814–1897). The mathematical model of the problem in high dimensions can be formulated as follows

$$\min\Big\{\max_{1\le i\le m}\|x-a_i\| \,:\, x\in\mathbb{R}^n\Big\}, \tag{2}$$

where $a_i\in\mathbb{R}^n$ for $i=1,\ldots,m$ are given points. Problem (2) is both a facility location problem and a major problem in computational geometry. The Sylvester problem and its versions in higher dimensions are also known under other names such as the smallest enclosing ball problem, the minimum ball problem, or the bomb problem. Over a century later, research on the smallest enclosing circle problem remains very active due to its important applications to clustering, nearest neighbor search, data classification, facility location, collision detection, computer graphics, and military operations. The problem has been widely treated in the literature from both theoretical and numerical standpoints.

In this dissertation, we use tools from nonsmooth analysis and optimization theory to study some complex facility location problems involving distances to sets in a finite dimensional space. In contrast to the existing facility location models where the locations are of negligible sizes, represented by points, the approach adapted in this dissertation allows us to deal with facility location problems where the locations are of non-negligible sizes, now represented by sets. Our efforts focus not only on studying theoretical aspects but also on developing effective solution methods for these problems.

The dissertation has five chapters, a list of references, and an appendix containing MATLAB codes of some numerical examples.

Chapter 1 collects several concepts and results from convex analysis and DC programming that are useful for subsequent studies. We also describe briefly the majorization-minimization principle, Nesterov's accelerated gradient method and smoothing technique, as well as P. D. Tao and L. T. H. An's DC algorithm.

Chapter 2 is devoted to numerically solving a number of new models of facility location which generalize the classical Fermat-Torricelli problem. Convergence of the proposed algorithms are proved and numerical tests are presented.

Chapter 3 studies a generalized version of problem (2) from both theoretical and numerical viewpoints. Sufficient conditions guaranteeing the existence and uniqueness of solutions, optimality conditions, constructions of the solutions in special cases are addressed. We also propose an algorithm based on the log-exponential smoothing technique and Nesterov's accelerated gradient method for solving the problem under consideration.

Chapter 4 is dedicated to studying a nonconvex facility location problem that is a generalization of problem (1). After establishing some theoretical properties, we propose an algorithm by combining the DC algorithm and the Weiszfeld algorithm for solving the problem.

Chapter 5 is totally different from the preceding parts of the dissertation. Motivated by some methods developed recently, we introduce a generalized proximal point algorithm for solving optimization problems in which the objective functions can be represented as differences of nonconvex and convex functions. Convergence of this algorithm under the main assumption that the objective function satisfies the Kurdyka-Łojasiewicz property is established.

Chapter 1

Preliminaries

Several concepts and results from convex analysis and DC programming are recalled in this chapter. As a preparation for the investigations in Chapters 2–5, we also describe the majorization-minimization principle, Nesterov's accelerated gradient method and smoothing technique, as well as DC algorithm.

1.1 Tools of Convex Analysis

We use $I\!R^n$ to denote the n-dimensional Euclidean space, $\langle\cdot,\cdot\rangle$ to denote the inner product, and $\|\cdot\|$ to denote the associated Euclidean norm. The *subdifferential* in the sense of convex analysis of a convex function $f : I\!R^n \to I\!R \cup \{+\infty\}$ at $\bar{x} \in \text{dom} f := \{x \in I\!R^n : \ f(x) < +\infty\}$ is defined by

$$\partial f(\bar{x}) := \{v \in I\!R^n : \ \langle v, x - \bar{x}\rangle \le f(x) - f(\bar{x}) \ \forall \ x \in I\!R^n\}.$$

For a nonempty closed convex subset Ω of $I\!R^n$ and a point $\bar{x} \in \Omega$, the *normal cone* to Ω at $\bar{x}$ is the set $N(\bar{x};\Omega) := \{v \in I\!R^n : \ \langle v, x - \bar{x}\rangle \le 0 \ \forall x \in \Omega\}$. This normal cone is the subdifferential of the *indicator function*

$$\delta(x;\Omega) = \begin{cases} 0 & \text{if } \ x \in \Omega, \\ +\infty & \text{if } \ x \notin \Omega, \end{cases}$$

at $\bar{x}$, i.e., $N(\bar{x};\Omega) = \partial\delta(\bar{x};\Omega)$. The *distance function* to Ω is defined by

$$d(x;\Omega) := \inf\{\|x - \omega\| : \ \omega \in \Omega\}, \quad x \in I\!R^n. \tag{1.1}$$

The notation $P(\bar{x};\Omega) := \{\bar{w} \in \Omega : \ d(\bar{x};\Omega) = \|\bar{x} - \bar{w}\|\}$ stands for the *Euclidean projection* from $\bar{x}$ to Ω. The subdifferential of the distance function (1.1) at $\bar{x}$ can be computed by the formula

$$\partial d(\bar{x};\Omega) = \begin{cases} N(\bar{x};\Omega) \cap I\!B & \text{if } \ \bar{x} \in \Omega, \\ \left\{\dfrac{\bar{x} - P(\bar{x};\Omega)}{d(\bar{x};\Omega)}\right\} & \text{if } \ \bar{x} \notin \Omega, \end{cases}$$

where $I\!B$ denotes the Euclidean closed unit ball of $I\!R^n$.

1.2 Majorization-Minimization Principle

The basic idea of majorization-minimization (MM) principle is to convert a hard optimization problem (for example, a non-differentiable problem) into a sequence of simpler ones (for example, smooth problems). The objective function $f: I\!R^n \to I\!R$ is said to be majorized by a surrogate function $\mathcal{M}: I\!R^n \times I\!R^n \to I\!R$ on Ω if $f(x) \le M(x,y)$ and $f(y) = \mathcal{M}(y,y)$ for all $x, y \in \Omega$. Given $x^0 \in \Omega$, the iterates of the associated MM algorithm for minimizing f on Ω are defined by

$$x^{k+1} \in \operatorname*{argmin}_{x\in\Omega} \mathcal{M}(x, x^k).$$

Because, $f(x^{k+1}) \le \mathcal{M}(x^{k+1}, x^k) \le \mathcal{M}(x^k, x^k) = f(x^k)$, the MM iterates generate a descent algorithm driving the objective function downhill.

1.3 Nesterov's Accelerated Gradient Method

Let $f: I\!R^n \to I\!R$ be a convex function with Lipschitz gradient. That is, there exists $\ell \ge 0$ such that $\|\nabla f(x) - \nabla f(y)\| \le \ell\|x-y\|$ for all $x, y \in I\!R^n$. Let Ω be a nonempty closed convex set. Yu. Nesterov (1983, 2005) considered the optimization problem

$$\min\big\{f(x):\ x\in\Omega\big\}. \tag{1.2}$$

Define $\Psi_\Omega(x) := \operatorname{argmin}\big\{\langle\nabla f(x), y-x\rangle + \frac{\ell}{2}\|x-y\|^2 :\ \ y\in\Omega\big\}$. Let d be a continuous and strongly convex function on Ω with modulus $\sigma > 0$. The function d is called a *prox-function* of the set Ω. Since d is a strongly convex function on the set Ω, it has a unique minimizer on this set. Denote $x^0 = \operatorname{argmin}\{d(x):\ x\in\Omega\}$. We can assume that $d(x^0) = 0$. Then Nesterov's accelerated gradient algorithm for solving (1.2) is outlined as follows.

INPUT: f, ℓ, $x^0 \in \Omega$
set $k = 0$
repeat
    find $y^k := \Psi_\Omega(x^k)$
    find $z^k := \operatorname{argmin}\big\{\frac{\ell}{\sigma}d(x) + \sum_{i=0}^{k}\frac{i+1}{2}[f(x^i) + \langle\nabla f(x^i), x - x^i\rangle]:\ x\in\Omega\big\}$
    set $x^k := \frac{2}{k+3}z^k + \frac{k+1}{k+3}y^k$
    set $k := k+1$
until a stopping criterion is satisfied.
OUTPUT: y^k.

1.4 Nesterov's Smoothing Technique

Let Ω be a nonempty closed convex subset of $I\!R^n$ and let Q be a nonempty compact convex subset of $I\!R^m$. Consider the constrained optimization problem (1.2) in which $f\colon \mathbb{R}^n \to \mathbb{R}$ is a convex function of the type

$$f(x) := \max\{\langle Ax, u\rangle - \phi(u) :\ u \in Q\},\ x \in \mathbb{R}^n,$$

where A is an $m\times n$ matrix and ϕ is a continuous convex function on Q. Let d_1 be a prox-function of Q with modulus $\sigma_1 > 0$ and $\bar{u} := \operatorname{argmin}\{d_1(u) :\ u \in Q\}$ be the unique minimizer of d_1 on Q. Assume that $d_1(\bar{u}) = 0$. We work mainly with $d_1(u) = \frac{1}{2}\|u - \bar{u}\|^2$ where $\bar{u} \in Q$. Let μ be a positive number called a *smooth parameter.* Define

$$f_\mu(x) := \max\{\langle Ax, u\rangle - \phi(u) - \mu d_1(u) :\ u \in Q\}. \tag{1.3}$$

Theorem 1.1 *The function f_μ in* (1.3) *is well defined and continuously differentiable on $\mathbb{R}^n$. The gradient of the function is $\nabla f_\mu(x) = A^\top u_\mu(x)$, where $u_\mu(x)$ is the unique element of Q such that the maximum in* (1.3) *is attained. Moreover, ∇f_μ is a Lipschitz function with the Lipschitz constant*

$$\ell_\mu = \frac{1}{\mu\sigma_1}\|A\|^2.$$

Let $D_1 := \max\{d_1(u) :\ u \in Q\}$. Then $f_\mu(x) \le f(x) \le f_\mu(x) + \mu D_1\ \forall x \in \mathbb{R}^n$.

1.5 DC Programming and DC Algorithm

Let $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $h : \mathbb{R}^n \to \mathbb{R}$ be convex functions. Here we assume that g is proper and lower semicontinuous. Consider the DC programming problem

$$\min\{f(x) := g(x) - h(x) :\ x \in \mathbb{R}^n\}. \tag{1.4}$$

Proposition 1.1 *If $\bar{x} \in \operatorname{dom} f$ is a local minimizer of* (1.4)*, then*

$$\partial h(\bar{x}) \subset \partial g(\bar{x}).$$

We use the convention $(+\infty) - (+\infty) = +\infty$. Toland's duality theorem can be stated as follows.

Proposition 1.2 *Under the assumptions made on the functions g and h, one has*

$$\inf\{g(x) - h(x) :\ x \in \mathbb{R}^n\} = \inf\{h^*(y) - g^*(y) :\ y \in \mathbb{R}^n\}.$$

The DCA for solving (1.4) is summarized as follows:

Step 1.	Choose $x^0 \in \operatorname{dom} g$.
Step 2.	For $k \ge 0$, use x^k to find $y^k \in \partial h(x^k)$.
	Then, use y^k to find $x^{k+1} \in \partial g^*(y^k)$.
Step 3.	Increase k by 1 and go back to **Step 2**.

Chapter 2

Effective Algorithms for Solving Generalized Fermat-Torricelli Problems

In this chapter, we present algorithms for solving a number of new models of facility location which generalize the classical Fermat-Torricelli problem. The chapter is written on the basis of the paper [2] in the list of author's related papers.

2.1 Generalized Fermat-Torricelli Problems

B. S. Modukhovich, N. M. Nam and J. Salinas (2012) proposed the following generalized model of the Fermat-Torricelli problem

$$\min\Big\{\mathcal{D}(x) := \sum_{i=1}^{m} d(x;\Omega_i) :\ x \in \Omega\Big\}, \tag{2.1}$$

where Ω and Ω_i for $i = 1, \dots, m$ are nonempty closed convex sets in $I\!R^n$ and

$$d(x;\Theta) := \inf\{\|x - w\| :\ w \in \Theta\} \tag{2.2}$$

is the Euclidean distance function to Θ. The authors mainly used the subgradient method for numerically solving (2.1). However, the subgradient method is known to be slow in general. Motivated by the question of finding better algorithms for solving (2.1), Eric. C. Chi and K. Lange (2014) proposed an algorithm that generalizes Weiszfeld's algorithm by invoking the majorization-minimization principle. We will follow the above research direction to deal with (2.1) when the distances under consideration are not necessarily Euclidean. The *generalized distance function* defined by the *dynamic* set F and the *target set* Θ is given by

$$d_F(x;\Theta) := \inf\{\sigma_F(x - w) :\ w \in \Theta\}, \tag{2.3}$$

where F is a nonempty compact convex set of $I\!R^n$ that contains the origin as an interior point. If F is the closed unit Euclidean ball of $I\!R^n$, the function (2.3) becomes the familiar distance function (2.2). We focus on developing algorithms for solving the following generalized version of (2.1)

$$\min\Big\{T(x) := \sum_{i=1}^{m} d_F(x;\Omega_i) : \ x \in \Omega\Big\}, \tag{2.4}$$

where Ω_i for $i = 1, \ldots, m$ and Ω are nonempty closed convex sets. The sets Ω_i for $i = 1, \ldots, m$ are called the *target sets* and the set Ω is called the *constraint set*. When all the target sets are singletons such as $\Omega_i = \{a_i\}$ for $i = 1, \ldots, m$, problem (2.4) reduces to

$$\min\Big\{H(x) := \sum_{i=1}^{m} \sigma_F(x - a_i) : \ x \in \Omega\Big\}. \tag{2.5}$$

Our approach can be outlined as follows. We first solve (2.5) by using Nesterov's smoothing techniques to approximate the nonsmooth function H by a smooth convex function with Lipschitz gradient. Then, the accelerated gradient methods are applied to the smooth problem. After that, we majorize the function T with a generalized version of MM principle and solve (2.4) by the MM algorithm. The convergence of the MM sequence is investigated under some appropriate assumptions.

2.2 Nesterov's Smoothing Technique and a General Form of the MM Principle

We now present a simplified version of Theorem 1.1 for which the gradient of f_μ has an explicit representation.

Theorem 2.1 *Let A be an $m \times n$ matrix and Q be a nonempty compact and convex subset of $\mathbb{R}^m$. Consider the function $f(x) := \max\{\langle Ax, u\rangle - \langle b, u\rangle : u \in Q\}$, $x \in \mathbb{R}^n$.*

Let $d(u) = \frac{1}{2}\|u - \bar{u}\|^2$ with $\bar{u} \in Q$. Then the function f_μ in (1.3) has the explicit representation

$$f_\mu(x) = \frac{\|Ax - b\|^2}{2\mu} + \langle Ax - b, \bar{u}\rangle - \frac{\mu}{2}\big[d(\bar{u} + \frac{Ax - b}{\mu}; Q)\big]^2$$

and is continuously differentiable on $\mathbb{R}^n$ with its gradient given by

$$\nabla f_\mu(x) = A^\top P(\bar{u} + \frac{Ax - b}{\mu}; Q).$$

The gradient ∇f_μ is a Lipschitz function with constant $\ell_\mu = \dfrac{1}{\mu}\|A\|^2$. Moreover, $f_\mu(x) \le f(x) \le f_\mu(x) + \dfrac{\mu}{2}[D(\bar{u}; Q)]^2$ for all $x \in \mathbb{R}^n$ with $D(\bar{u}; Q) := \sup\{\|\bar{u} - u\| : \ u \in Q\}$.

We continue with a more general version of MM principle. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and let Ω be a nonempty closed convex subset of $\mathbb{R}^n$. Consider the optimization problem

$$\min\{f(x) : x \in \Omega\}. \tag{2.6}$$

Let $\mathcal{M} : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ and let $\mathcal{F} : \mathbb{R}^n \rightrightarrows \mathbb{R}^p$ be a set-valued mapping with nonempty values such that the following properties hold for all $x, y \in \mathbb{R}^n$:

$$f(x) \le \mathcal{M}(x, z) \ \ \forall z \in \mathcal{F}(y), \text{ and } f(x) = \mathcal{M}(x, z) \ \ \forall z \in \mathcal{F}(x).$$

Given $x^0 \in \Omega$, the MM algorithm to solve (2.6) is given by

$$\text{Choose } z^k \in \mathcal{F}(x^k) \text{ and find } x^{k+1} \in \operatorname{argmin}\{\mathcal{M}(x, z^k) : x \in \Omega\}.$$

2.3 Problems Involving Points

We say that F is *normally smooth* if for every boundary point x of F there exists $a_x \in \mathbb{R}^n$ such that $N(x; F)$ is the cone generated by a_x. Let $\mathbb{B}^*_F := \{u \in \mathbb{R}^n : \sigma_F(u) \le 1\}$.

Proposition 2.1 *F is normally smooth if and only if $\mathbb{B}^*_F$ is strictly convex.*

Proposition 2.2 *Suppose that F is normally smooth. If for any $x, y \in \Omega$ with $x \ne y$, the line connecting x and y, $\mathcal{L}(x, y)$, does not contain at least one of the points a_i for $i = 1, \ldots, m$, then problem (2.5) has a unique optimal solution.*

Given any $\bar{u} \in F$, consider the smooth approximation function given by

$$H_\mu(x) := \sum_{i=1}^{m} \left(\frac{\|x - a_i\|^2}{2\mu} + \langle x - a_i, \bar{u} \rangle - \frac{\mu}{2}[d(\bar{u} + \frac{x - a_i}{\mu}; F)]^2 \right). \tag{2.7}$$

Proposition 2.3 *The function H_μ defined by (2.7) is continuously differentiable on $\mathbb{R}^n$ with its gradient given by*

$$\nabla H_\mu(x) = \sum_{i=1}^{m} P(\bar{u} + \frac{x - a_i}{\mu}; F).$$

The gradient ∇H_μ is a Lipschitz function with constant $\mathcal{L}_\mu = \dfrac{m}{\mu}$. Moreover, one has the following estimate

$$H_\mu(x) \le H(x) \le H_\mu(x) + m\frac{\mu}{2}[D(\bar{u}; F)]^2 \quad \forall x \in \mathbb{R}^n.$$

We are now ready to write a pseudocode for solving problem (2.5).

INPUT: a_i for $i = 1, \ldots, m$, μ.
INITIALIZE: Choose $x^0 \in \Omega$ and set $\ell = \dfrac{m}{\mu}$.
Set $k = 0$
Repeat the following
Compute $\nabla H_\mu(x^k) = \sum_{i=1}^m P(\bar{u} + \dfrac{x^k - a_i}{\mu}; F)$.
Find $y^k := P(x^k - \frac{1}{\ell}\nabla H_\mu(x^k); \Omega)$.
Find $z^k := P(x^0 - \frac{1}{\ell}\sum_{i=0}^k \dfrac{i+1}{2}\nabla H_\mu(x^i); \Omega)$.
Set $x^{k+1} := \dfrac{2}{k+3} z^k + \dfrac{k+1}{k+3} y^k$.
until a stopping criterion is satisfied.
OUTPUT: y^k.

2.4 Problems Involving Sets

The *generalized projection* from a point $x \in I\!R^n$ to a set Θ is defined by $\pi_F(x;\Theta) := \{w \in \Theta : \sigma_F(x - w) = d_F(x;\Theta)\}$. A convex set F is said to be *normally round* if $N(x;F) \neq N(y;F)$ for any distinct boundary points x, y of F.

Proposition 2.4 *Given a nonempty closed convex set Θ, consider the generalized distance function* (2.3)*. Then the following properties hold:*
(i) $|d_F(x;\Theta) - d_F(y;\Theta)| \leq \|F\|\,\|x - y\|$ for all $x, y \in I\!R^n$.
(ii) *The function $d_F(\cdot;\Theta)$ is convex, and $\partial d_F(\bar{x};\Theta) = \partial\sigma_F(\bar{x} - \bar{w}) \cap N(\bar{w};\Theta)$ for any $\bar{x} \in I\!R^n$, where $\bar{w} \in \pi_F(\bar{x};\Theta)$ and this representation does not depend on the choice of $\bar{w}$.*
(iii) *If F is normally smooth and round, then $\sigma_F(\cdot)$ is differentiable at any nonzero point, and $d_F(\cdot;\Theta)$ is continuously differentiable on the complement of Θ with $\nabla d_F(\bar{x};\Theta) = \nabla\sigma_F(\bar{x} - \bar{w})$, where $\bar{x} \notin \Theta$ and $\bar{w} := \pi_F(\bar{x};\Theta)$.*

Proposition 2.5 *Suppose that F is normally smooth and the target sets Ω_i for $i = 1, \ldots, m$ are strictly convex with at least one of them being bounded. If for any $x, y \in \Omega$, with $x \neq y$, there exists an index $i \in \{1, \ldots, m\}$ such that $\pi_F(x;\Omega_i) \notin \mathcal{L}(x,y)$. Then problem* (2.4) *has a unique optimal solution.*

Let us apply the MM principle to the generalized Fermat-Torricelli problem. We rely on the following properties which hold for all $x, y \in I\!R^n$:

(i) $d_F(x;\Theta) = \sigma_F(x - w)$ for all $w \in \pi_F(x;\Theta)$.

(ii) $d_F(x;\Theta) \leq \sigma_F(x - w)$ for all $w \in \pi_F(y;\Theta)$.

Consider the set-valued mapping $\mathcal{F}(x) := \Pi_{i=1}^m \pi_F(x;\Omega_i)$. Then the cost function $T(x)$ is majorized by

$$T(x) \leq \mathcal{M}(x,w) := \sum_{i=1}^m \sigma_F(x - w^i),\ w = (w^1, \ldots, w^m) \in \mathcal{F}(y).$$

Moreover, $T(x) = \mathcal{M}(x, w)$ whenever $w \in \mathcal{F}(x)$. Thus, given $x^0 \in \Omega$, the MM iteration is given by

$$x^{k+1} \in \operatorname{argmin}\{\mathcal{M}(x, w^k) : \ x \in \Omega\} \ \text{ with } \ w^k \in \mathcal{F}(x^k).$$

This algorithm can be written more explicitly as follows.

INPUT: Ω and m target sets Ω_i, $i = 1, \ldots, m$.
INITIALIZE: $x^0 \in \Omega$.
Set $k = 0$.
Repeat the following
Find $y^{k,i} \in \pi_F(x^k; \Omega_i)$.
Solve the following problem with a stopping criterion
$\min_{x\in\Omega} \sum_{i=1}^m \sigma_F(x - y^{k,i})$,
and denote its solution by x^{k+1}.
until a stopping criterion is satisfied.
OUTPUT: x^k.

Proposition 2.6 *Consider the generalized Fermat-Torricelli problem* (2.4) *in which F is normally smooth and round. Let $\{x^k\}$ be the sequence in the MM algorithm defined by* $x^{k+1} \in \operatorname{argmin}\left\{\sum_{i=1}^m \sigma_F(x - \pi_F(x^k; \Omega_i)) : \ x \in \Omega\right\}$. *Suppose that $\{x^k\}$ converges to $\bar{x}$ that does not belong to Ω_i for $i = 1, \ldots, m$. Then $\bar{x}$ is an optimal solution of problem* (2.4).

Lemma 2.1 *Consider the generalized Fermat-Torricelli problem* (2.4) *in which at least one of the target sets Ω_i for $i = 1, \ldots, m$ is bounded and F is normally smooth and round. Suppose that the constraint set Ω does not intersect any of the target sets Ω_i for $i = 1, \ldots, m$, and for any $x, y \in \Omega$ with $x \neq y$ the line connecting x and y, $\mathcal{L}(x, y)$, does not intersect at least one of the target sets. For any $x \in \Omega$, consider the mapping $\psi : \Omega \to \Omega$ defined by*

$$\psi(x) := \operatorname{argmin}\left\{\sum_{i=1}^m \sigma_F(y - \pi_F(x; \Omega_i)) : \ y \in \Omega\right\}.$$

Then ψ is continuous at any point $\bar{x} \in \Omega$, and $T(\psi(x)) < T(x)$ whenever $x \neq \psi(x)$.

Theorem 2.2 *Consider problem* (2.4) *in the setting of Lemma 2.1. Let $\{x^k\}$ be a sequence generated by the MM algorithm, i.e., $x^{k+1} = \psi(x^k)$ with a given $x^0 \in \Omega$. Then any cluster point of the sequence $\{x^k\}$ is an optimal solution of problem* (2.4). *If we assume additionally that Ω_i for $i = 1, \ldots, m$ are strictly convex, then $\{x^k\}$ converges to the unique optimal solution of the problem.*

It is important to note that the algorithm may not converge in general. Our examples (given in the dissertation) partially answer the question raised by E. Chi, H. Zhou, and K. Lange (2013).

Chapter 3

The Smallest Intersecting Ball Problem

We study the following generalized version of the smallest enclosing circle problem: *Given a finite number of nonempty closed convex sets in* $\mathbb{R}^n$*, find a ball with the smallest radius that intersects all of the sets.* After establishing many theoretical properties, based on the log-exponential smoothing technique and Nesterov's accelerated gradient method, we present an effective algorithm for solving this problem. This chapter is written on the basis of the papers [1] and [4].

3.1 Problem Formulation and Theoretical Aspects

Given a set $P = \{p_1, \ldots, p_m\} \subset \mathbb{R}^n$. The *smallest enclosing ball problem* (SEBP, for brevity) asks for the ball of smallest radius that contains P. This problem can be formulated as

$$\min\left\{\max_{1\le i\le m} ||x - p_i|| : \; x \in \mathbb{R}^n\right\}. \tag{3.1}$$

Let Ω_i for $i = 1, \ldots, m$ and Ω be nonempty closed convex subsets of $\mathbb{R}^n$. For any $x \in \Omega$, there always exists $r > 0$ such that

$$\mathbb{B}(x; r) \cap \Omega_i \neq \emptyset \quad \text{for all} \quad i = 1, \ldots, m. \tag{3.2}$$

The *smallest intersecting ball problem* (SIBP, for brevity) generated by the target sets Ω_i for $i = 1, \ldots, m$ and the constraint set Ω, asks for a ball with the smallest radius $r > 0$ (if exists) that satisfies property (3.2). Consider the optimization problem

$$\min\left\{\mathcal{D}(x) := \max_{1\le i\le m} d(x; \Omega_i) : \; x \in \Omega\right\}. \tag{3.3}$$

When $\Omega = \mathbb{R}^n$, we have the unconstrained problem

$$\min\left\{\mathcal{D}(x) : \; x \in \mathbb{R}^n\right\}. \tag{3.4}$$

We use the standing assumption: $\bigcap_{i=1}^n (\Omega_i \cap \Omega) = \emptyset$. The following result allows us to identify SIBP with problem (3.3).

Proposition 3.1 *Consider problem* (3.3)*. Then* $\bar{x} \in \Omega$ *is an optimal solution of this problem with* $r = \mathcal{D}(\bar{x})$ *if and only if* $I\!B(\bar{x}; r)$ *is a smallest ball that satisfies* (3.2)*.*

Proposition 3.2 *Suppose that at least one of the sets* $\Omega, \Omega_1, \ldots, \Omega_m$ *is bounded. Then the smallest intersecting ball problem* (3.3) *has a solution.*

Theorem 3.1 *Suppose that the target sets* Ω_i*, for* $i = 1, \ldots, m$*, are strictly convex, and at least one of the sets among* $\Omega, \Omega_1, ..., \Omega_m$ *is bounded. Then the smallest intersecting ball problem* (3.3) *has a unique optimal solution if and only if* $\bigcap_{i=1}^{m} (\Omega \cap \Omega_i)$ *contains at most one point.*

For each $x \in \Omega$, the set of active indices for $\mathcal{D}$ at x is defined by

$$I(x) = \{i \in \{1, \ldots, m\} : \mathcal{D}(x) = d(x; \Omega_i)\}.$$

Proposition 3.3 *A point* $\bar{x} \in \Omega$ *is an optimal solution of problem* (3.3) *if and only if*

$$\bar{x} \in \mathrm{co}\{\bar{\omega}_i : \ i \in I(\bar{x})\} - N(\bar{x}; \Omega),$$

where $\bar{\omega}_i = P(\bar{x}; \Omega_i)$ *and* $\mathrm{co}M$ *denotes the convex hull of a subset* $M \subset \mathbb{R}^n$*.*

Corollary 3.1 *A point* $\bar{x}$ *is a solution of problem* (3.4) *if and only if*

$$\bar{x} \in \mathrm{co}\{\bar{\omega}_i : i \in I(\bar{x})\},$$

where $\bar{\omega}_i = P(\bar{x}; \Omega_i)$*. In particular, if* $\Omega_i = \{a_i\}$*,* $i = 1, \ldots, m$*, then* $\bar{x}$ *is the solution of* (3.1) *generated by* a_i*,* $i = 1, \ldots, m$*, if and only if*

$$\bar{x} \in \mathrm{co}\{a_i : \ i \in I(\bar{x})\}.$$

It is obvious that $\mathrm{co}\{a_i : i \in I(\bar{x})\} \subset \mathrm{co}\{a_i : i = 1, \ldots, m\}$. Thus our result in Corollary 3.1 covers Theorem 3.6 in the paper of L. Drager, J. Lee and C. Martin (2007).

We also show that a smallest intersecting ball generated by m convex sets in $I\!R^n$ can be determined by at most $n + 1$ sets among them.

Proposition 3.4 *Consider problem* (3.4) *in which* Ω_i*,* $i = 1, \ldots, m$*, are disjoint. Suppose that* $I\!B(\bar{x}; r)$ *is a smallest intersecting ball of the problem. Then there exists an index set* J *with* $2 \leq |J| \leq n + 1$ *such that* $I\!B(\bar{x}; r)$ *is also a smallest intersecting ball of* (3.4) *in which the target sets are* Ω_j*,* $j \in J$*.*

The next result is a generalization of Theorem 4.4 in the paper of L. Drager, J. Lee and C. Martin (2007).

Theorem 3.2 *Consider the smallest intersecting ball problem* (3.4) *generated by the closed balls* $\Omega_i = I\!B(\omega_i; r_i)$*,* $i = 1, \ldots, m$*. Let* $r_{\min} = \min_{1 \leq i \leq m} r_i$*,*

$r_{\max} := \max_{1\le i\le m} r_i$, $\ell = \min\{n+1, m\}$, $P = \{\omega_i : i = 1,\ldots,m\}$ *and let* $I\!B(\bar{x}; r)$ *be the smallest intersecting ball. Then*

$$\frac{1}{2}\text{diam}(P) - r_{\max} \le r \le \sqrt{\frac{\ell-1}{2\ell}}\text{diam}(P) - r_{\min}.$$

where $\text{diam}(P) := \max\{\|x-y\| : x, y \in P\}$.

3.2 A Smoothing Technique for SIBP

For $p > 0$, the *log-exponential smoothing function* of $\mathcal{D}$ is defined by

$$\mathcal{D}(x,p) = p \ln \sum_{i=1}^{m} \exp\left(\frac{G_i(x,p)}{p}\right), \tag{3.5}$$

where $G_i(x,p) := \sqrt{d(x;\Omega_i)^2 + p^2}$. The sets Ω_i for $i = 1,\ldots,m$ are said to be *non-collinear* if it is impossible to draw a straight line that intersects all of these sets.

Theorem 3.3 *The function* $\mathcal{D}(x,p)$ *defined in* (3.5) *has the following properties:*

(i) *If* $x \in I\!R^n$ *and* $0 < p_1 < p_2$, *then* $\mathcal{D}(x,p_1) < \mathcal{D}(x,p_2)$.
(ii) *For any* $x \in I\!R^n$ *and* $p > 0$, $0 \le \mathcal{D}(x,p) - \mathcal{D}(x) \le p(1 + \ln m)$.
(iii) *For any* $p > 0$, *the function* $\mathcal{D}(\cdot,p)$ *is convex. If we suppose further that the sets* Ω_i *for* $i = 1,\ldots,m$ *are strictly convex and non-collinear, then* $\mathcal{D}(\cdot,p)$ *is strictly convex.*
(iv) *For any* $p > 0$, $\mathcal{D}(\cdot,p)$ *is continuously differentiable with the gradient in* x *computed by*

$$\nabla_x \mathcal{D}(x,p) = \sum_{i=1}^{m} \frac{\Lambda_i(x,p)}{G_i(x,p)}\left(x - \widetilde{x}_i\right),$$

where $\widetilde{x}_i := P(x;\Omega_i)$, *and*

$$\Lambda_i(x,p) := \frac{\exp\left(G_i(x,p)/p\right)}{\sum_{i=1}^{m}\exp\left(G_i(x,p)/p\right)}.$$

(v) *If at least one of the target sets* Ω_i *for* $i = 1,\ldots,m$ *is bounded, then* $\mathcal{D}(\cdot,p)$ *is coercive in the sense that* $\lim_{\|x\|\to+\infty}\mathcal{D}(x,p) = +\infty$.

3.3 A MM Algorithm for SIBP

Proposition 3.5 *Let* $\{p_k\}$ *be a sequence of positive real numbers converging to* 0. *For each* k, *let* $y^k \in \text{argmin}_{x\in\Omega}\mathcal{D}(x,p_k)$. *Then* $\{y^k\}$ *is a bounded* *sequence and every cluster point of* $\{y^k\}$ *is an optimal solution of* (3.3)*. Suppose further that* (3.3) *has a unique optimal solution. Then* $\{y^k\}$ *converges to that optimal solution.*

For $x, y \in I\!R^n$ and $p > 0$, define

$$\mathcal{G}(x, y, p) := p \ln \sum_{i=1}^{m} \exp\left(\frac{\sqrt{\|x - P(y; \Omega_i)\|^2 + p^2}}{p}\right).$$

Choose a small number $\bar{p} > 0$. In order to solve (3.3), we solve the problem

$$\min \{\mathcal{D}(x, \bar{p}) : x \in \Omega\} \tag{3.6}$$

by using the MM algorithm.

Proposition 3.6 *Given* $\bar{p} > 0$ *and* $x^0 \in \Omega$*, the sequence* $\{x^k\}$ *defined by*

$$x^k := \text{argmin}_{x\in\Omega}\, \mathcal{G}(x, x^{k-1}, \bar{p}),$$

has a convergent subsequence.

The convergence of the MM algorithm depends on the algorithm map:

$$\psi(x) := \underset{y\in\Omega}{\text{argmin}}\, \mathcal{G}(y, x, \bar{p}). \tag{3.7}$$

Theorem 3.4 *Given* $\bar{p} > 0$*, the function* $\mathcal{D}(\cdot, \bar{p})$ *and the algorithm map* $\psi : \Omega \to \Omega$ *defined by* (3.7) *satisfy the following conditions:*

(i) *For* $x^0 \in \Omega$*, the set* $\mathcal{L}(x^0) := \{x \in \Omega : \mathcal{D}(x, \bar{p}) \le \mathcal{D}(x^0, \bar{p})\}$ *is compact.*
(ii) ψ *is continuous on* Ω*.*
(iii) $\mathcal{D}(\psi(x), \bar{p}) < \mathcal{D}(x, \bar{p})$ *whenever* $x \neq \psi(x)$*.*
(iv) *Any fixed point* $\bar{x}$ *of* ψ *is a minimizer of* $\mathcal{D}(\cdot, \bar{p})$ *on* Ω*.*

Corollary 3.2 *Given* $\bar{p} > 0$ *and* $x^0 \in \Omega$*, the sequence* $\{x^k\}$ *with* $x^k := \text{argmin}_{x\in\Omega}\, \mathcal{G}(x, x^{k-1}, \bar{p})$ *has a subsequence that converges to an optimal solution of* (3.6)*. If we suppose further that problem* (3.6) *has a unique optimal solution, then* $\{x^k\}$ *converges to this optimal solution.*

It has been experimentally observed that, in order to get a more effective algorithm, instead of choosing a small value p ahead of time, we decrease its value gradually. Our algorithm is outlined as follows.

```
INPUT: Ω, p_0 > 0, x^0 ∈ Ω, m target sets Ω_i, i = 1,...,m, N, σ ∈ (0,1)
set p = p_0
for k = 1,...,N do
    use Nesterov's accelerated gradient method to solve the following problem
        x^k := argmin_{x∈Ω} G(x, x^{k-1}, p)
    until a stopping criterion is satisfied
    set p := σp
end for
OUTPUT: x^N
```

Chapter 4

A Nonconvex Location Problem Involving Sets

This chapter is devoted to study a location problem that involves a weighted sum of distances to closed convex sets. As several of the weights might be negative, traditional solution methods of convex optimization are not applicable. After obtaining some existence theorems, we introduce a simple algorithm for solving the problem. Our method is based on the Pham Dinh - Le Thi algorithm for DC programming and a generalized version of the Weiszfeld algorithm, which works well for convex location problems. This chapter is written on the basis of the paper [3].

4.1 Problem Formulation

We will be concerned with the following constrained optimization problem

$$\min \left\{ f(x) := \sum_{i=1}^{p} \alpha_i d(x; \Omega_i) - \sum_{j=1}^{q} \beta_j d(x; \Theta_j) : \ x \in S \right\}, \tag{4.1}$$

where $\{\Omega_i : \ i = 1, \ldots, p\}$ and $\{\Theta_j : \ j = 1, \ldots, q\}$ are two finite collections of nonempty closed convex sets in $I\!R^n$, S is a nonempty closed convex constraint set and the real numbers α_i and β_j are all positive.

4.2 Solution Existence in the General Case

Define $I = \{1, \ldots, p\}, \quad J = \{1, \ldots, q\}$. The following result generalizes Theorem 1 in the paper of Z. Drezner and G. O. Wesolowsky (1990).

Theorem 4.1 (Sufficient conditions for the solution existence) *Problem* (4.1) *has a solution if at least one of the following conditions is satisfied:*

(i) *S is bounded;*
(ii) $\sum\limits_{i \in I} \alpha_i > \sum\limits_{j \in J} \beta_j$, *and all the sets Ω_i, $i \in I$, are bounded.*

Proposition 4.1 *If $\sum_{i \in I} \alpha_i < \sum_{j \in J} \beta_j$, S is unbounded, and all the sets Θ_j, $j \in J$, are bounded, then* $\inf\{f(x) : \ x \in S\} = -\infty$; *so* (4.1) *has no solution.*

Proposition 4.2 *If* $\sum_{i\in I}\alpha_i = \sum_{j\in J}\beta_j$*, and all of the sets* Ω_i*,* $i \in I$ *and* Θ_j*,* $j \in J$*, are bounded, then there exists* $\gamma > 0$ *such that* $|f(x)| \leq \gamma$ *for all* $x \in I\!R^n$*.*

If the equality

$$\sum_{i\in I}\alpha_i = \sum_{j\in J}\beta_j, \tag{4.2}$$

holds, then the solution set of (4.1) may be nonempty or empty as well. We now provide a sufficient condition for the solution existence under the assumption (4.2).

Proposition 4.3 *Any solution of the problem*

$$\max\left\{h(x) := \sum_{j\in J}\beta_j d(x;\Theta_j)\;:\; x \in \Omega_1\right\}$$

is a solution of (4.1) *in the case where* $\Omega_1 \subset S$*,* $I = \{1\}$*, and* $\alpha_1 = \sum_{j\in J}\beta_j$*. Thus, in that case, if* Ω_1 *is bounded then* (4.1) *has a solution.*

Sufficient conditions forcing the solution set of (4.1) to be a subset of one of the sets Ω_i are given in the next proposition, which is an extension of the Proposition 3 in P.-C. Chen, P. Hansen, B. Jaumard, and H. Tuy (1992).

Proposition 4.4 *Consider problem* (4.1) *where* $\Omega_{i_0} \subset S$ *for some* $i_0 \in I$*, and*

$$\alpha_{i_0} > \sum_{i\in I\setminus\{i_0\}}\alpha_i + \sum_{j\in J}\beta_j.$$

Then any solution of (4.1) *must belong to* Ω_{i_0}*.*

To show that (4.1) can have an empty solution set under condition (4.2), let us consider a special case where $S = I\!R^n$, $\Omega_i = \{a_i\}$, $\Theta_j = \{b_j\}$ with a_i and b_j, $i \in I$ and $j \in J$, being some given points. Problem (4.1) now becomes

$$\min\left\{f(x) = \sum_{i\in I}\alpha_i\|x - a_i\| - \sum_{j\in J}\beta_j\|x - b_j\|\;:\; x \in I\!R^n\right\}. \tag{4.3}$$

The next lemma regarding the value of the cost function at infinity is a generalization of Lemma 4 in the paper of Z. Drezner and G. O. Wesolowsky (1990).

Lemma 4.1 *Let* $f(x)$ *be given as in* (4.3) *and let* $w = \sum_{i\in I}\alpha_i a_i - \sum_{j\in J}\beta_j b_j$*. If* $w = 0$*, then*

$$\lim_{\|x\|\to+\infty} f(x) = 0.$$

If $w \neq 0$*, then*

$$\liminf_{\|x\|\to+\infty} f(x) = -\|w\|.$$

Proposition 4.5 *Let $I = \{1, \ldots, p\}, p \geq 2$, and let $b \in I\!R^n$. If $\beta = \sum_{i \in I} \alpha_i$ and the vectors $\{a_i - b\}$ for $i \in I$ are linearly independent, then the problem*

$$\min \left\{ f(x) = \sum_{i \in I} \alpha_i \|x - a_i\| - \beta \|x - b\| : \ x \in I\!R^n \right\}$$

has no solution.

4.3 Solution Existence in a Special Case

Consider a special case of problem (4.1) where $p = q = 1$ and $S = I\!R^n$; that is,

$$\min \{ f(x) := \alpha d(x; \Omega) - \beta d(x; \Theta) : \ x \in I\!R^n \}, \tag{4.4}$$

where $\alpha \geq \beta > 0$. We are going to establish several properties of the optimal solutions to problem (4.4). The relationship between (4.4) and the problem

$$\max\{d(x; \Theta) : \ x \in \Omega\} \tag{4.5}$$

will also be discussed.

Proposition 4.6 *If $\alpha > \beta$, then $\bar{x}$ is a solution of* (4.4) *if and only if it is a solution of* (4.5). *Thus, in the case $\alpha > \beta$, the solution set of* (4.4) *does not depend on the choice of α and β.*

We now describe a relationship between the solution sets of (4.4) and (4.5), which are denoted respectively by S_1 and S_2.

Proposition 4.7 *Suppose that $\Omega \setminus \Theta \neq \emptyset$. If $\alpha = \beta$, then*

$$S_1 = \{ \bar{u} + I\!R^+ (\bar{u} - P(\bar{u}; \Theta)) : \ \bar{u} \in S_2 \}.$$

4.4 A Combination of DCA and Generalized Weiszfeld Algorithm

To solve (4.1) by the DCA, we rewrite (4.1) equivalently as

$$\min \{ g(x) - h(x) : \ x \in I\!R^n \}.$$

where

$$g(x) := \sum_{i \in I} \alpha_i d(x; \Omega_i) + \frac{\lambda}{2}\|x\|^2 + \delta(x; S), \ h(x) := \sum_{j \in J} \beta_j d(x; \Theta_j) + \frac{\lambda}{2}\|x\|^2,$$

and $\lambda > 0$ being an arbitrarily chosen constant. An element $y^k \in \partial h(x^k)$ can be chosen by $y^k = \sum_{j \in J} u^{k,j} + \lambda x^k$, where

$$u^{k,j} = \begin{cases} \beta_j \dfrac{x^k - P(x^k;\Theta_j)}{d(x^k;\Theta_j)}, & \text{if } x^k \notin \Theta_j, \\ 0, & \text{otherwise.} \end{cases} \tag{4.6}$$

To find $x^{k+1} \in \partial g^*(y^k)$, we solve the following problems by Weiszfeld's algorithm

$$(P_v) \qquad \min \left\{ \varphi_v(x) := \sum_{i \in I} \alpha_i d(x;\Omega_i) + \frac{\lambda}{2}\|x\|^2 - \langle v, x \rangle \, : \, x \in S \right\}.$$

For simplicity, assume that $\Omega_i \cap S = \emptyset$ for every $i \in I$. Define the mapping

$$F_v(x) = \frac{\displaystyle\sum_{i \in I} \frac{\alpha_i P(x;\Omega_i)}{d(x;\Omega_i)} + v}{\displaystyle\sum_{i \in I} \frac{\alpha_i}{d(x;\Omega_i)} + \lambda}, \quad x \in S. \tag{4.7}$$

We introduce the following *generalized Weiszfeld algorithm* to solve (P_v):

- Choose $x^0 \in S$.
- Find $x^{k+1} = P(F_v(x^k); S)$ for $k \in \mathbb{N}$, where F_v is defined in (4.7).

Theorem 4.2 *Consider the generalized Weiszfeld algorithm for solving* (P_v). *If* $x^{k+1} \neq x^k$, *then* $\varphi_v(x^{k+1}) < \varphi_v(x^k)$.

Theorem 4.3 *The sequence* $\{x^k\}$ *produced by the generalized Weiszfeld algorithm converges to the unique solution of problem* (P_v).

Combining the DCA and the generalized Weiszfeld algorithm, we get the following algorithm for solving (4.1).

```
INPUT: x^0 ∈ S, λ > 0, Ω_i for i = 1,...,p and Θ_j for j = 1,...,q.
set k = 0
for k = 1,...,N do
    Find y^k according to (4.6)
    Find the unique solution x^{k+1} = argmin_{x∈S} φ_{y^k}(x) by the generalized Weiszfeld algorithm
    provided that a stopping criterion and a starting point z_k are given.
    set k := k + 1
end for
OUTPUT: x^{N+1}.
```

Theorem 4.4 *Consider the above algorithm for solving* (4.1). *If either condition* (i) *or* (ii) *in Theorem 4.1 is satisfied, then any limit point of the iterative sequence* $\{x^k\}$ *is a critical point of* (4.1).

Chapter 5

Convergence Analysis of a Proximal Point Algorithm for Minimizing Differences of Functions

In this chapter, we introduce a generalized proximal point algorithm to minimize the difference of a nonconvex function and a convex function. We also study convergence results of this algorithm under the main assumption that the objective function satisfies the Kurdyka - Łojasiewicz property. This chapter is written on the basis of the paper [5].

5.1 The Kurdyka-Łojasiewicz Property

For a lower semicontinuous function $f : I\!R^n \to I\!R \cup \{+\infty\}$ with $\bar{x} \in \text{dom} f$, the *Fréchet subdifferential* of f at $\bar{x}$ is defined by

$$\partial^F f(\bar{x}) = \left\{ v \in I\!R^n : \liminf_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - \langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \geq 0 \right\}.$$

We set $\partial^F f(\bar{x}) = \emptyset$ if $\bar{x} \notin \text{dom} f$. Based on the Fréchet subdifferential, the *limiting/Mordukhovich subdifferential* of f at $\bar{x} \in \text{dom} f$ is defined by

$$\partial^L f(\bar{x}) = \operatorname*{Limsup}_{x \xrightarrow{f} \bar{x}} \partial^F f(x) = \{ v \in I\!R^n : \exists\, x^k \xrightarrow{f} \bar{x}, v^k \in \partial^F f(x^k), v^k \to v \},$$

where the notation $x \xrightarrow{f} \bar{x}$ means that $x \to \bar{x}$ and $f(x) \to f(\bar{x})$. We also set $\partial^L f(\bar{x}) = \emptyset$ if $\bar{x} \notin \text{dom} f$. The Clarke subdifferential of a locally Lipschitz continuous function f at $\bar{x}$ can be represented via the limiting subdifferential as $\partial^C f(\bar{x}) = \text{co}\, \partial^L f(\bar{x})$.

Following H. Attouch, J. Bolte, P. Redont, and A. Soubeyran (2010), a lower semicontinuous function $f \colon I\!R^n \to I\!R \cup \{+\infty\}$ satisfies the *Kurdyka - Łojasiewicz property* (KŁ property) at $x^* \in \text{dom}\, \partial^L f$ if there exist $\eta > 0$, a neighborhood U of x^*, and a continuous concave function $\varphi : [0, \eta) \to [0, +\infty)$ with (i) $\varphi(0) = 0$, (ii) φ is of class C^1 on $(0, \eta)$, (iii) $\varphi' > 0$ on $(0, \eta)$, and (iv) for every $x \in U$ with $f(x^*) < f(x) < f(x^*) + \eta$, we have

$$\varphi'\left(f(x) - f(x^*)\right) \text{dist}\left(0, \partial^L f(x)\right) \geq 1. \tag{5.1}$$

We say that f satisfies the *strong Kurdyka - Łojasiewicz* property at x^* if the same assertion holds for the Clarke subdifferential $\partial^C f(x)$. According to H. Attouch, J. Bolte, P. Redont, and A. Soubeyran (2010), for a proper lower semicontinuous function $f : I\!R^n \to I\!R \cup \{+\infty\}$, the Kurdyka - Łojasiewicz property is satisfied at any point $\bar{x} \in \mathrm{dom}\partial^L f$ such that $0 \notin \partial^L f(\bar{x})$. A subset Ω of $I\!R^n$ is called *semi-algebraic* if it can be represented as a finite union of sets of the form $\{x \in I\!R^n : \ p_i(x) = 0, q_i(x) < 0 \ \text{ for all } i = 1, \dots, m\}$, where p_i and q_i for $i = 1, \dots, m$ are polynomial functions. A function f is said to be semi-algebraic if its graph $\{(x; y) \in I\!R^{n+1} : \ y = f(x)\}$, is a semi-algebraic subset of $I\!R^{n+1}$. It is known that a proper lower semicontinuous semi-algebraic function $f : I\!R^n \to I\!R \cup \{+\infty\}$ satisfies the Kurdyka - Łojasiewicz property at all points in $\mathrm{dom}\, \partial^L f$ with $\varphi(s) = cs^{1-\theta}$ for some $\theta \in [0, 1)$ and $c > 0$.

5.2 A Generalized Proximal Point Algorithm for Minimizing a Difference of Functions

We now focus on the convergence analysis of a proximal point algorithm for solving nonconvex optimization problems of the type

$$\min \{f(x) = g_1(x) + g_2(x) - h(x) : \ x \in I\!R^n\}, \tag{5.2}$$

where $g_1(x)\colon I\!R^n \to I\!R \cup \{+\infty\}$ is proper and lower semicontinuous, $g_2(x) : I\!R^n \to I\!R$ is differentiable with L - Lipschitz gradient, and $h : I\!R^n \to I\!R$ is convex. The specific structure of (5.2) is flexible enough to include the problem of minimizing a smooth function on a closed constraint set $\min\{g(x) : x \in \Omega\}$, and the general DC problem:

$$\min \{f(x) = g(x) - h(x) : x \in I\!R^n\}, \tag{5.3}$$

where $g\colon I\!R^n \to I\!R \cup \{+\infty\}$ is a proper lower semicontinuous convex function and $h\colon I\!R^n \to I\!R$ is convex.

Proposition 5.1 *If $\bar{x} \in \mathrm{dom}\, f$ is a local minimizer of the function f considered in* (5.2)*, then*

$$\partial h(\bar{x}) \subset \partial^L g_1(\bar{x}) + \nabla g_2(\bar{x}). \tag{5.4}$$

Any point $\bar{x} \in \mathrm{dom} f$ satisfying condition (5.4) is called a *stationary point* of (5.2). In general, this condition is hard to be reached and we may relax it to $[\partial^L g_1(\bar{x}) + \nabla g_2(\bar{x})] \cap \partial h(\bar{x}) \neq \emptyset$ and call $\bar{x}$ a *critical point* of f. Let $g : I\!R^n \to I\!R \cup \{+\infty\}$ be a proper lower semicontinuous function. The *Moreau proximal mapping*, with regularization parameter $t > 0$, is defined by

$$\mathrm{prox}_t^g(x) = \mathrm{argmin}\left\{g(u) + \frac{t}{2}\|u - x\|^2 : \ u \in I\!R^n\right\}.$$

Generalized Proximal Point Algorithm (GPPA)

INPUT: f, $x^0 \in \text{dom}\, g_1$ and $t > L$
set $k = 0$
repeat
find $y^k \in \partial h(x^k)$.
find x^{k+1} as follows

$$x^{k+1} \in \text{prox}_t^{g_1}\left(x^k - \frac{\nabla g_2(x^k) - y^k}{t}\right).$$

set $k := k + 1$
until a stopping criterion is satisfied.
OUTPUT: x^k

Theorem 5.1 *Consider the* GPPA *for solving* (5.2) *in which* $g_1(x)\colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is proper and lower semicontinuous with* $\inf_{x\in\mathbb{R}^n} g_1(x) > -\infty$, $g_2(x)\colon \mathbb{R}^n \to \mathbb{R}$ *is differentiable with L - Lipschitz gradient, and* $h : \mathbb{R}^n \to \mathbb{R}$ *is convex. Then*
(i) *For any* $k \geq 1$, *we have* $f(x^k) - f(x^{k+1}) \geq \frac{t-L}{2}\|x^k - x^{k+1}\|^2$.
(ii) *If* $\alpha = \inf\limits_{x\in\mathbb{R}^n} f(x) > -\infty$, *then* $\lim\limits_{k\to+\infty} f(x^k) = \ell^* \geq \alpha$, $\lim\limits_{k\to+\infty} \|x^k - x^{k+1}\| = 0$.
(iii) *If* $\alpha = \inf\limits_{x\in\mathbb{R}^n} f(x) > -\infty$ *and* $\{x^k\}$ *is bounded, then every cluster point of* $\{x^k\}$ *is a critical point of* f.

Proposition 5.2 *Suppose that* $\inf_{x\in\mathbb{R}^n} f(x) > -\infty$, f *is proper and lower semicontinuous. If the* GPPA *sequence* $\{x^k\}$ *has a cluster point* x^*, *then* $\lim\limits_{k\to+\infty} f(x^k) = f(x^*)$. *Thus,* f *has the same value at all cluster points of* $\{x^k\}$.

The forthcoming theorems establish sufficient conditions that guarantee the convergence of the sequence $\{x^k\}$ generated by the GPPA. Let C^* denote the set of cluster points of the sequence $\{x^k\}$.

Theorem 5.2 *Suppose that* $g_1(x)\colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is proper and lower semicontinuous with* $\inf_{x\in\mathbb{R}^n} g_1(x) > -\infty$, $g_2(x)\colon \mathbb{R}^n \to \mathbb{R}$ *is differentiable with L - Lipschitz gradient, and* $h : \mathbb{R}^n \to \mathbb{R}$ *is convex. Suppose further that* ∇h *is* $L(h)$ *- Lipschitz continuous,* $\inf_{x\in\mathbb{R}^n} f(x) > -\infty$, *and* f *has the Kurdyka - Łojasiewicz property at any point* $x \in \text{dom} f$. *If* $C^* \neq \emptyset$, *then the* GPPA *sequence* $\{x^k\}$ *converges to a critical point of* f.

In the next result, we assume that $g_1(x) = 0$ and put $g_2(x) = g(x)$.

Theorem 5.3 *Let* $f = g - h$ *with* $\inf_{x\in\mathbb{R}^n} f(x) > -\infty$. *Suppose that* g *is differentiable and* ∇g *is L - Lipschitz continuous,* f *has the strong Kurdyka - Łojasiewicz property at any point* $x \in \text{dom} f$, *and* h *is a finite convex function. If* $C^* \neq \emptyset$, *then the* GPPA *sequence* $\{x^k\}$ *converges to a critical point of* f.

General Conclusions

This dissertation has applied variational analysis and optimization theory to complex facility location problems involving distances to sets. In contrast to the existing facility location models where the locations are of negligible sizes, represented by points, the new approach allows us to deal with facility location problems where the locations are of non-negligible sizes, now represented by sets. Our efforts focused not only on studying theoretical aspects but also on developing effective algorithms for solving these problems. Besides, we also introduced an algorithm for minimizing the difference of functions.

Our main results include:

- Algorithms based on Nesterov's smoothing technique and the majorization-minimization principle for solving new models of the Fermat-Torricelli problem.

- Theoretical properties as well as an algorithm based on the log-exponential smoothing technique and Nesterov's accelerated gradient method for the smallest intersecting ball problem.

- Solution existence together with an algorithm based on the DC algorithm and the Weiszfeld algorithm for nonconvex facility location problems.

- Convergence analysis of a generalized proximal point algorithm for minimizing the difference of a nonconvex function and a convex function.

The techniques used in the dissertation are not only applicable to single facility location problems but also open up the possibility of applications to other fields such as multi-facility location problems, split feasibility problems, support vector machines, image processing. These are interesting topics for our future research.

List of Author's Related Papers

1. N. T. An, D. Giles, N. M. Nam, and R. B. Rector, *The log-exponential smoothing technique and Nesterov's accelerated gradient method for generalized Sylvester problems*, J. Optim. Theory. Appl., **168** (2016), No 2, 559–583.
2. N. T. An and N. M. Nam, *Convergence analysis of a proximal point algorithm for minimizing differences of functions*, to appear in Optimization.
3. N. T. An, N. M. Nam, and N. D. Yen, *A D.C. algorithm via convex analysis approach for solving a location problem involving sets*, J. Convex Anal., **23** (2016), No. 1, 77–101.
4. N. M. Nam, N. T. An, R. B. Rector, and J. Sun, *Nonsmooth algorithms and Nesterov's smoothing technique for generalized Fermat-Torricelli problems*, SIAM J. Optim., **24** (2014), No. 4, 1815–1839.
5. N. M. Nam, N. T. An, and J. Salinas, *Applications of convex analysis to the smallest intersecting ball problem*, J. Convex Anal., **19** (2012), No. 2, 497–518.